

Снижение размерности гиперспектральных данных с использованием нелинейных автокодировщиков

Е.В. Мясников^{1,2}

¹Самарский национальный исследовательский университет им. академика С.П. Королева, Московское шоссе 34А, Самара, Россия, 443086

²Институт систем обработки изображений РАН - филиал ФНИЦ «Кристаллография и фотоника» РАН, Молодогвардейская 151, Самара, Россия, 443001

Аннотация. Известной особенностью гиперспектральных изображений является высокое спектральное разрешение, которое позволяет идентифицировать материалы и классифицировать объекты на изображениях с высокой точностью. Однако гиперспектральные изображения содержат существенную избыточность, которая может быть устранена с помощью методов снижения размерности. В этой работе мы предлагаем и изучаем два метода снижения размерности, основанных на предварительном обучении нейронной сети кодер-декодер результатами методов нелинейного отображения и анализа главных компонент. Эксперименты, выполненные на открытом наборе данных, показывают, что предложенные методы обеспечивают возможность получения признаков с хорошей разделяющей способностью и позволяют реконструировать исходные гиперспектральные данные с малой ошибкой.

1. Введение

В настоящее время гиперспектральные изображения широко используются в различных областях, таких как сельское хозяйство, медицина, биология, химия и т.д. Известной особенностью гиперспектральных изображений является высокое спектральное разрешение, которое позволяет идентифицировать материалы и классифицировать изображенные на сцене объекты с высокой точностью.

Однако гиперспектральные изображения содержат существенную избыточность, которая может быть устранена с помощью методов снижения размерности. Изображения, полученные после этапа снижения размерности, могут обрабатываться более эффективно, так как при обработке задействуется гораздо меньший объем данных. Ключевым требованием к процедурам снижения размерности является способность сохранять качество решения прикладных задач, таких как классификация, сегментация, обнаружение материала и т.д.

Хотя существует ряд процедур снижения размерности общего назначения [1], их использование в анализе гиперспектральных изображений ограничено, поскольку многие из них выполняют только прямое отображение данных и не обеспечивают возможности восстановления исходных гиперспектральных данных.

В последние годы нейросетевые подходы становятся все более и более популярными. В частности, для снижения размерности гиперспектральных изображений были использованы автокодировщики [2]. Такие нейронные сети выполняют как нелинейное снижение

размерности, так и обеспечивают обратное отображение, которое позволяет восстановить исходные гиперспектральные данные с точностью до некоторой ошибки восстановления.

Недавно было показано [3], что автокодировщики могут быть предварительно обучены с использованием метода главных компонент, а их использование для снижения размерности позволило превзойти метод PCA как с точки зрения ошибки восстановления, так и точности классификации.

Однако было также показано [4], что метод нелинейного отображения [5] имеет преимущества перед PCA с точки зрения качества классификации и сегментации гиперспектральных изображений. По этой причине в настоящей работе мы изучаем возможность обучения автокодировщиков для выполнения нелинейного отображения [5]. В частности, мы разделяем автокодировщик на кодер и декодер и обучаем обе части отдельно, используя результаты нелинейного отображения, а также оцениваем эффект последующей тонкой настройки всей сети.

Работа имеет следующую структуру. В следующем разделе 2 дается необходимая теоретическая информация об архитектуре нейронной сети и методе нелинейного отображения. В разделе 3 описывается процедура обучения и результаты экспериментов. Статья завершается заключением и списком литературы приведены в конце статьи.

2. Предлагаемый метод

2.1. Автокодировщик

Нейронная сеть автокодировщик, предложенная в [2] и называемая первоначально автоассоциативной нейронной сетью, состоит из двух последовательных частей, кодера и декодера.

Кодер принимает на вход многомерный вектор $x \in R^M$ в качестве входных данных и создает соответствующие представления $y \in R^m$ малой размерности $m < M$. Кодер состоит как минимум из двух полносвязных слоев. Первый слой содержит некоторое количество нейронов (определяемое параметрами архитектуры нейронной сети), связанных со всеми компонентами входного вектора. Последний слой кодера содержит количество нейронов, равное требуемой размерности редуцированного пространства.

Декодер обычно имеет зеркально отраженную архитектуру. Он имеет такое же количество слоев с тем же количеством нейронов, хотя это не является обязательным требованием. В любом случае, входной слой декодера принимает на вход редуцированное представление $y \in R^m$ с выхода кодера и восстанавливает многомерные векторы $\tilde{x} \in R^M$. Таким образом, выходной слой декодера имеет количество нейронов, равное входной размерности M . Количество скрытых слоев и нейронов определяется параметрами архитектуры нейронной сети.

Поскольку число нейронов в выходном слое кодера меньше, чем число нейронов во входном и скрытом слоях, этот слой часто называют бутылочным горлышком, и архитектура сети также носит подобное название.

Сети с архитектурой автокодировщика обычно обучаются в режиме самообучения путем подачи одних и тех же многомерных векторов $x \in R^M$ на входной и выходной слои автокодировщика. Сам процесс обучения основан на минимизации следующей функции потерь:

$$E = \frac{1}{N} \sum_{i=1}^N \|x_i - \tilde{x}_i\|^2 \quad (1)$$

где N - размер выборки, а $x \in R^M$ и $\tilde{x} \in R^M$ - входы и выходы сети, соответственно. После обучения кодер может использоваться для снижения размерности исходных данных (прямое отображение), а декодер может использоваться для восстановления исходных данных по их редуцированному представлению (обратное отображение).

В настоящей работе мы изучаем, могут ли кодер и декодер обучаться отдельно, чтобы заставить нейронную сеть выполнить отображение с желаемыми свойствами. Ранее было показано, что отдельное предварительное обучение кодера и декодера с использованием

результатов PCA помогло выполнить обучение более эффективно по сравнению со стандартным обучением.

В частности, подход, предложенный в [3], состоит из следующих этапов: выполнение PCA для входного набора данных; предварительное обучение кодера формированию результатов PCA по входным данным; предварительное обучение декодера восстановлению исходных данных по их кодированному представлению; тонкая настройка всей сети по стандартной схеме.

В настоящей работе мы следуем аналогичной схеме, но используем результаты метода нелинейного отображения вместо PCA, чтобы понять, может ли такой подход быть более эффективным, чем стандартный PCA, нелинейное отображение или автокодировщик предварительно обученный с использованием PCA [3].

2.2. Нелинейное отображение

Нелинейное отображение - это численная процедура, которая выполняет отображение (нефункциональное) данных в редуцированное пространство таким образом, чтобы сохранить структуру данных (см., например, [5]). При этом структура определяется всеми попарными расстояниями между точками в наборе данных. Для измерения расстояний обычно используется Евклидово расстояние $d()$.

Поскольку в общем случае попарные расстояния не могут быть сохранены точно, вводится так называемая ошибка отображения данных:

$$\varepsilon = \mu \cdot \sum_{i,j=1(i < j)}^N \left(\rho_{i,j} (d(x_i, x_j) - d(y_i, y_j))^2 \right). \quad (2)$$

Здесь N - количество точек, $d(x_i, x_j)$ - расстояние между точками x_i и x_j в многомерном пространстве, $d(y_i, y_j)$ - расстояние между соответствующими точками y_i, y_j в редуцированном пространстве, μ и ρ некоторые константы. Обычно μ - это инверсия суммы квадратов расстояний между всеми возможными парами точек данных в многомерном пространстве, а ρ_{ij} равны единице.

Минимизация ошибки отображения данных обычно выполняется с использованием метода градиентного спуска. Настраиваемыми параметрами являются координаты точек данных $y_i \in R_m$.

В этой работе для минимизации ошибки отображения данных мы используем стохастический градиентный спуск, основанный на малых случайных подвыборках (mini batch). В целом, метод снижения размерности с использованием нелинейного отображения состоит в инициализации координат y_i результатами анализа главных компонент с последующим уточнением y_i с использованием стохастического градиентного спуска. Процесс оптимизации (уточнения) останавливается после стабилизации координат точек данных y_j в редуцированном пространстве.

2.3. Используемые в исследовании методы

Как было указано во введении, в этой статье мы изучаем несколько вариантов обучения сети с архитектурой автокодировщика. В частности, рассматриваются следующие методы:

- автокодировщик, предварительно обученный с использованием результатов PCA (AE-PCA), как описано в [3];
- нейронная сеть кодер-декодер, части которой обучаются по-отдельности с использованием результатов метода нелинейного отображения (ED-NLM);
- та же самая сеть автокодировщика, предварительно обученная с использованием результатов метода нелинейного отображения и настроенная с использованием стандартного подхода (AE-NLM).

3. Эксперименты

В этом разделе описываются результаты экспериментов, выполненных с использованием набора данных Indian Pines. Этот набор данных был получен с использованием

гиперспектрального сенсора AVIRIS. Изображение содержит 145 x 145 пикселей и 224 спектральных компоненты [6]. Из-за высокого уровня шума и поглощения воды на исходном изображении мы использовали версию, содержащую 204 спектральные компоненты.

Во всех описанных экспериментах для реализации нейронных сетей мы использовали библиотеку Keras и язык Python. Эксперименты проводились с использованием графического процессора GeForce GTX 1070 ti.

Для каждого рассматриваемого нейросетевого метода мы варьировали количество скрытых слоев в кодере и декодере и проводили эксперименты для одного и двух скрытых слоев, что соответствует четырех- и шестислойному автокодировщику.

Количество нейронов во входном слое кодера и в выходном слое декодера определялось размерностью входного пространства, равной числу каналов в гиперспектральном изображении. Количество нейронов в среднем слое варьировалось от 1 до 10 в зависимости от размерности редуцированного пространства. Количество нейронов в скрытых слоях также варьировалось. В частности, в скрытых слоях использовалось 64, 128 и 256 нейронов.

В соответствии с рекомендациями, приведенными в [3], мы использовали функции активации ReLU для скрытых слоев и линейные функции активации в выходных слоях кодера и декодера. Аналогично, использовался оптимизатор Adam [7] с параметрами по умолчанию. Размер подвыборки был равен 16, однако, мы предполагаем, что возможно использовать и больший размер подвыборки.

Чтобы измерить эффективность каждого конкретного подхода, мы оценивали как ошибку восстановления, определенную в (1), так и точность классификации при использовании редуцированного представления.

В качестве последнего индикатора использовалась общая точность классификатора по одному ближайшему соседу (1-NN). Сама точность измерялась как доля правильно классифицированных пикселей изображения в общем числе классифицируемых пикселей. Для измерения точности сначала выполнялось снижение размерности для всех пикселей рассматриваемого изображения с использованием одного из исследуемых методов. Затем все пиксели с известной истинной классификацией разделялись на обучающую и тестовую выборки в пропорции 60/40. После этого с использованием обучающего множества производилось обучение классификатора, и, с использованием тестовой выборки, оценивалась его точность.

В первом эксперименте разные методы и архитектуры, описанные в подразделе 2.3, сравнивались с точки зрения ошибки реконструкции (1). Результаты этого эксперимента показаны на рис. 1. В частности, кодер и декодер сети AE-PCA предварительно обучались в течение 50 итераций. Далее, в течение еще 50 итераций, настраивалась сеть целиком, после чего оценивалось качество восстановления.

Сеть AE-NLM обучалась с использованием той же стратегии. Отличие заключалось в использовании результатов NLM вместо результатов PCA на этапе предварительного обучения. Для сети ED-NLM кодер и декодер обучались в течение 100 эпох. После обучения в качестве показателя качества оценивалась ошибка (1). Эксперименты проводились для разного количества слоев и нейронов.

Как видно из рисунка, ошибка восстановления уменьшается с ростом размерности m редуцированного пространства, определяемой числом нейронов в среднем слое, что является ожидаемым результатом.

Хотя мы не можем выделить какой-либо метод как победитель в этом эксперименте, мы должны отметить, что метод ED-NLM достаточно часто показывает лучшие результаты. Это означает, что результат нелинейного отображения, который использовался для обучения, дает возможность восстановить исходные данные с довольно хорошим качеством. Это также означает, что декодер, обученный на данных NLM, может использоваться как обратное отображение для NLM.

Во втором эксперименте мы сравнили рассмотренные методы с точки зрения точности классификации. Результаты этого эксперимента показаны на рисунке 2. На этом рисунке мы

добавили результаты для классического линейного (PCA) и нелинейного (NLM) методов снижения размерности.

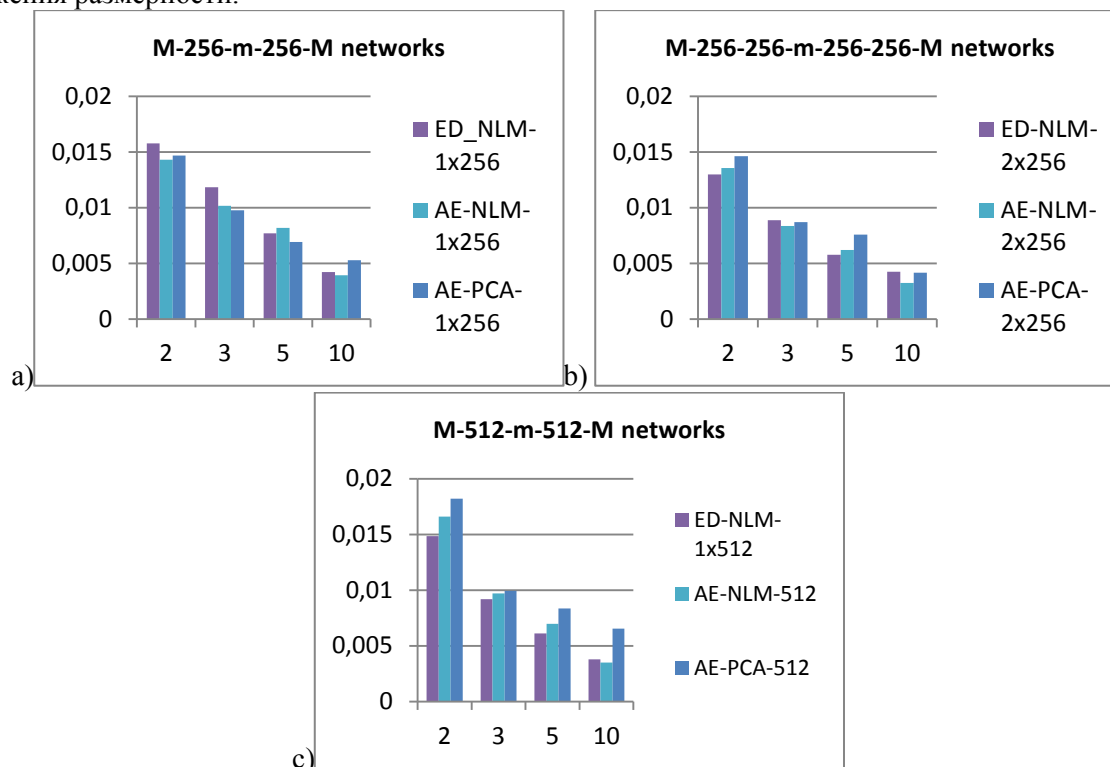


Рисунок 1. Зависимость ошибки восстановления (по вертикальной оси) от размерности редуцированного пространства (по горизонтальной оси) для различных методов и сетевых архитектур (а-с).

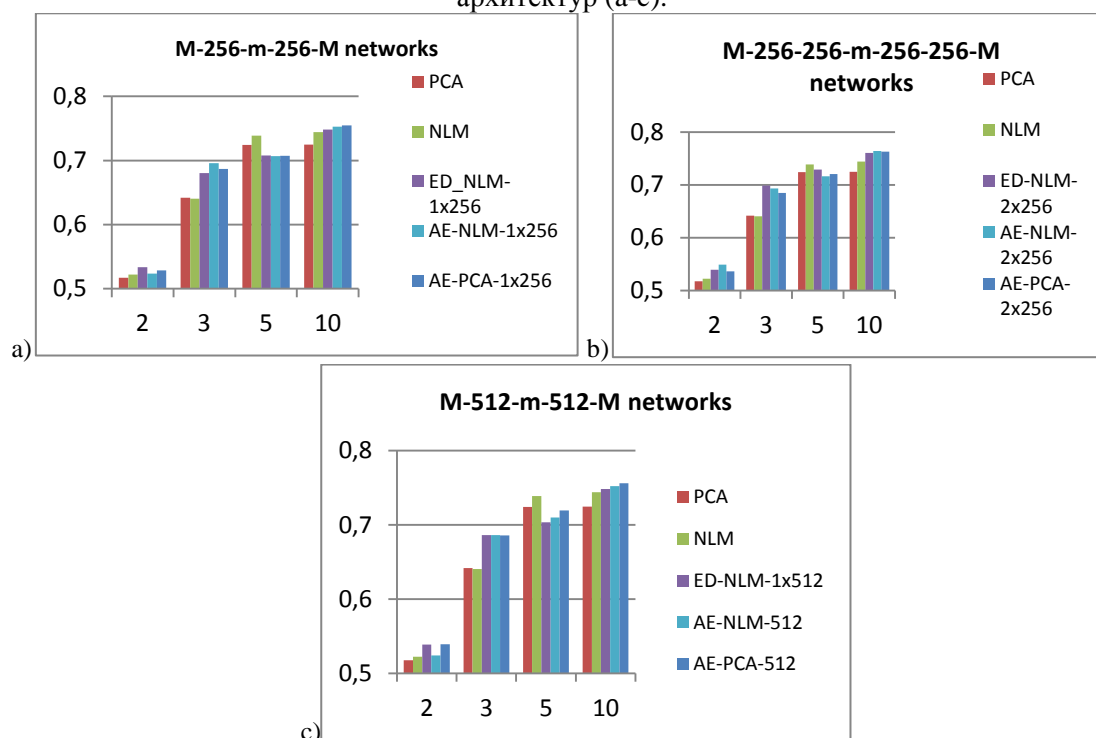


Рисунок 2. Зависимость точности классификации (по вертикальной оси) от размерности редуцированного пространства (по горизонтальной оси) для различных методов и сетевых архитектур (а-с).

Как видно, методы, основанные на автокодировщиках, в большинстве случаев дали лучшие результаты, чем классические подходы. Однако выделить среди них какой-либо один подход не представляется возможным. Можно также отметить, что не наблюдаем каких-либо существенных преимуществ в тонкой настройке инициализированной сети NLM (AE-PCA) по сравнению с версией с отдельным кодером и декодером (ED-PCA).

4. Заключение

В настоящей работе мы изучили несколько нейросетевых методов снижения размерности, основанных на архитектуре автокодировщика. Мы сравнили предложенные методы как с точки зрения ошибки восстановления, так и точности попиксельной классификации. Мы показали, что в большинстве случаев предложенные методы превосходят базовые (PCA и NLM), как с точки зрения ошибки восстановления, так и точности классификации. Декодер, обученный с использованием результатов NLM, может быть успешно использован для выполнения обратного отображения при анализе гиперспектральных изображений.

5. Благодарности

Работа выполнена при финансовой поддержке РФФИ в рамках научного проекта № 18-07-01312-а в частях «2. Метод» - «3. Эксперименты и результаты» и Министерства науки и высшего образования РФ в рамках госзадания ФНИЦ «Кристаллография и фотоника» РАН в частях «1. Введение» и «4. Заключение».

6. Литература

- [1] Lee, J.A. Nonlinear Dimensionality Reduction / J.A. Lee, M. Verleysen - Springer, 2007.
- [2] Kramer, M.A. Nonlinear principal component analysis using autoassociative neural networks / M.A. Kramer // AIChE J. – 1991. – Vol. 37. – P. 233-243.
- [3] Myasnikov, E. Dimensionality Reduction of Hyperspectral Images using Autoassociative Neural Networks / E. Myasnikov // IEEE Proc. of International Multi-Conference on Engineering, Computer and Information Sciences (SIBIRCON). – 2019. – P. 0591-0595.
- [4] Myasnikov, E. Evaluation of nonlinear dimensionality reduction techniques for classification of hyperspectral images / E. Myasnikov // CEUR Workshop Proceedings. – 2018. – Vol. 2268. – P. 147-154.
- [5] Sammon, J.W. A nonlinear mapping for data structure analysis / J.W. Sammon // IEEE Transactions on Computers. – 1969. – Vol. 18(5). – P. 401-409.
- [6] Baumgardner, M.F. 220 Band AVIRIS Hyperspectral Image Data Set: June 12, 1992 Indian Pine Test Site 3 / M.F. Baumgardner, L.L. Biehl, D.A. Landgrebe // Purdue University Research Repository, 2015. DOI:10.4231/R7RX991C.
- [7] Kingma, D. Adam: A Method for Stochastic Optimization / D. Kingma, J. Ba // arXiv: 1412.6980v8, 2017.

Hyperspectral data dimensionality reduction using nonlinear autoencoders

E.V. Myasnikov^{1,2}

¹Samara National Research University, Moskovskoe Shosse 34A, Samara, Russia, 443086

²Image Processing Systems Institute of RAS - Branch of the FSRC "Crystallography and Photonics" RAS, Molodogvardejskaya street 151, Samara, Russia, 443001

Abstract. The known feature of hyperspectral images is a high spectral resolution, which allows to identify materials and classify objects in images with high accuracy. However hyperspectral images contain substantial redundancy, which can be eliminated with the aid of dimensionality reduction techniques. In this paper, we propose and study several dimensionality reduction techniques based on the pretraining the encoder-decoder neural network with the results of the nonlinear mapping and principal component analysis techniques. The experiments performed on an open dataset show that the proposed techniques both provide the discriminative low-dimensional features and allow to reconstruct source hyperspectral data with little error.